

Detecting unusual time trends: A Bayesian mixture modelling approach

G. Li¹, N. Best¹, P. Clark² and S. Richardson¹

28 August, 2009

1. Department of Epidemiology and Public Health, Imperial College
London, UK
2. Office for National Statistics, UK.

Royal Geographical Society Annual International Conference, Manchester

Outline

- 1 Motivation: Analysis of space-time data, $Y_{i,t}$
- 2 A Bayesian Mixture model
- 3 Application: Modelling Income
- 4 Simulation: Disease Surveillance
- 5 Conclusion and Discussion

Income modelling

- The temporal trend patterns of average household income are likely to be similar across different administrative regions (e.g., counties and municipalities).
- However, some regions may display unexpected changes over time due to, for example, different responses to policies and/or social changes.
- Detection of areas with “unusual” time trends provides a tool to
 - evaluate impacts of policies;
 - assist allocating resources efficiently to areas that exhibit “downward” trends.

Disease surveillance

- Incidence of a disease varies over time and across space though the variations are often seen to be small.
- Abrupt changes in time could manifest some short-latency risk factors.
- To detect areas with "unusual" changes over time is of great importance to both epidemiological explanation and, perhaps more importantly, to effective intervention.

Some issues ...

- Modelling income
 - A typical survey is constructed to sample only some percentage (e.g., 1%) of the population
 - Local trend estimations may not be reliable due to variability created by small samples.
- Disease surveillance
 - Sparseness of the count data when disaggregated over time and space.

A Bayesian hierarchical mixture model

- We have proposed a Bayesian mixture modelling framework that
 - estimates the common temporal trend;
 - captures substantive changes of the local temporal patterns;
 - hence facilitates the detection of unusual time trends.
- We will demonstrate the performance of our mixture model using
 - 1 an income dataset from the Swedish LOUISE population registry;
 - 2 a simulation study mimicking disease surveillance.

- 1 Motivation
- 2 Mixture model
- 3 Application: Modelling Income
 - Data
 - Results
 - Summary
- 4 Simulation: Disease Surveillance
 - Data and model
 - Results
 - Summary
- 5 Conclusion and Discussion

Model specification

- We employ the Bayesian hierarchical modelling approach.
- For simplicity, we describe the model using the income application.
- Denoting $\hat{Y}_{i,t}$ and $\hat{\sigma}_{i,t}^2$ the average income per household and the corresponding design variance of municipality i at time t obtained from a typical 1% sample.

- First level

$$\hat{Y}_{i,t} \sim N(\mu_{i,t}, \hat{\sigma}_{i,t}^2) \quad (1)$$

- Second level

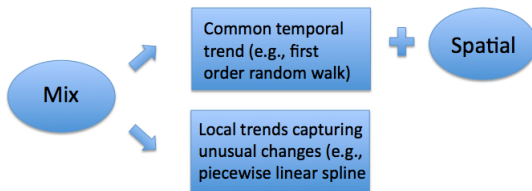
$$\mu_{i,t} = \alpha + \beta \cdot \bar{X}_{i,t} + \text{mix}_{i,t} \quad (2)$$

where $\bar{X}_{i,t}$ are some area level covariates.

Spec. for the mixture term

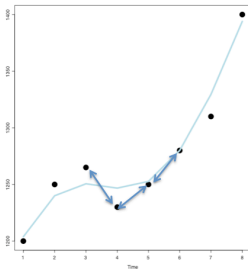
- For fitting a local trend, we use either a trend that is common for all areas or a trend that is specific to that area.
- For the **common trend** model, we use a **first order random walk**.
- For the **area-specific** model, we use a **piecewise linear spline model (PWLS)**.
- We introduce a latent allocation variable z_i for the selection:

$$\text{mix}_{i,t}|z_i = \begin{cases} RW_t + \eta_i & \text{if } z_i = 1 \text{ (common)} \\ \text{PWLS}_{i,t} & \text{if } z_i = 0 \text{ (area-spec.)} \end{cases} \quad (3)$$

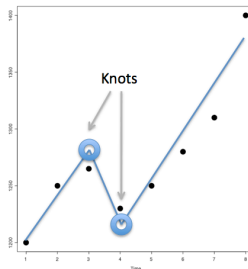


Illustrations of RW and PWLS

- For RW, estimation for one time point borrows information from the "neighboring" time points so that the resulting common trend is smooth, as we would expect.
- For PWLS, a straight line is fitted between two knots and two straight lines are joint "continuously" so that it can capture large/abrupt changes.



First order random walk (RW)



Piecewise Linear Spline

A computation issue

- Model fitting is done in WinBUGS, a software for doing Bayesian analysis.
- In particular fitting of the local splines is done using the reversible jump interface in WinBUGS, which allows us to treat the number of knots, the positions of the knots and other spline coefficients. (This is difficult!)
- As a result, standard implementation of mixture models did not work because the model always selected the "easy fitting" RW model.
- Inspired by Carlin and Chib (1995), we implemented an algorithm that runs both trend models in parallel.
- We can then select one or the other at each iteration.
- This implementation speeds up convergence!

- 1 Motivation
- 2 Mixture model
- 3 Application: Modelling Income**
 - Data
 - Results
 - Summary
- 4 Simulation: Disease Surveillance
 - Data and model
 - Results
 - Summary
- 5 Conclusion and Discussion

Data

- The data set is from the LOUISE Population registry in Sweden from 1992 to 1999.
- It contains various socio-economic variables at both the individual and household levels.
- The variable of interest is the average equivalised income per household for each of the 284 municipalities in the country.
- We will NOT use the entire data set.
- Instead we use data from a mock survey in which we sample data on income for only **1%** of the total number of households in each municipality in each year.
- 50 replicates of the 1% sample were simulated.
- The number of HHs in the 1% data ranges from 14 to 307.

Assessing performance

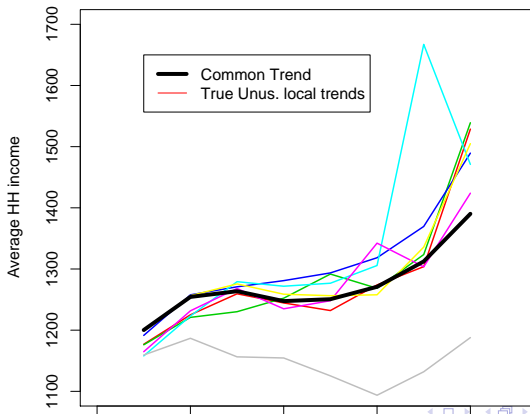
- To assess the detection performance, we compare the results from fitting the detection model to the "mock" survey data to some "true" unusual trends defined using the entire dataset.

Defining the "true" unusual trends from the entire data

- To define a "true" unusual time trend, we need to decide
 - 1 the number of time points at which the local trend deviates from the common trend;
 - 2 whether it is unusual
 - 3 the amount it differing from the common trend at each time point;

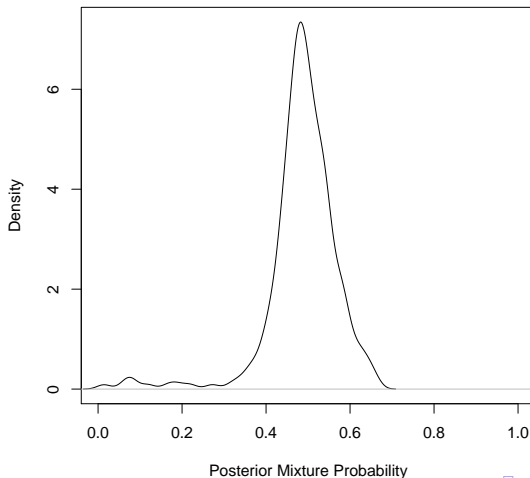
Illustration of true unusual trends (3pts + medium)

There are 30 areas classified as unusual in this case using the entire data.

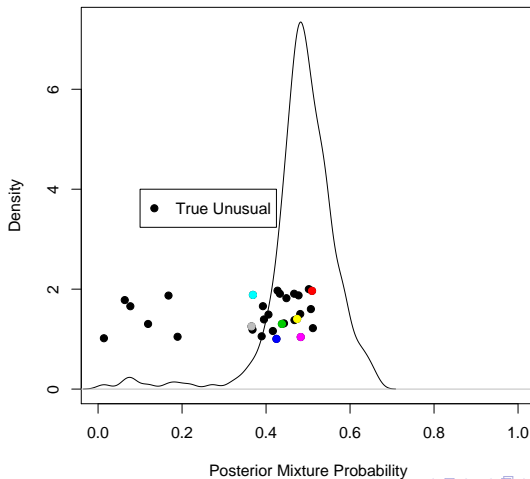


- Results from fitting the detection model to the 1% survey data

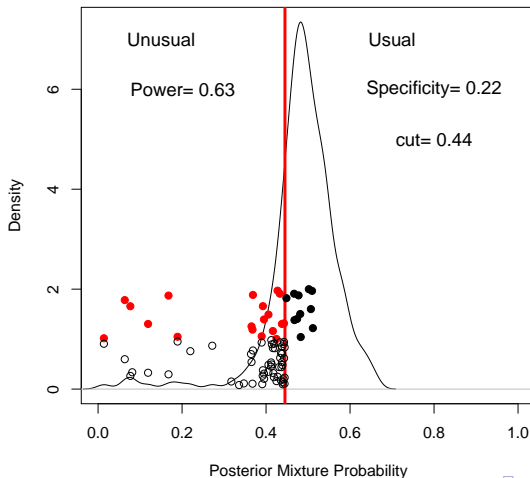
Histogram of the probabilities being usual (284 areas)



Where the true unusual trends are



Classification of unusual areas based on the probs



Receiver Operating Characteristic (ROC) curve

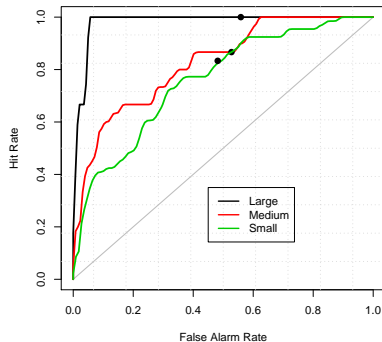


Figure: The "True" local trends differing from the common trend at 3 time points

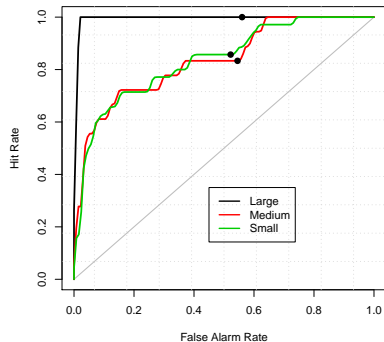


Figure: The "True" local trends differing from the common trend at 5 time points

Area under a ROC curve (AUC)

Magnitude of the difference

Number of departing time points (tps)

AUC	Small	Medium	Large
3 tps	0.76	0.82	0.98
4 tps	0.81	0.81	0.98
5 tps	0.84	0.84	0.99
6 tps	0.83	0.89	0.99

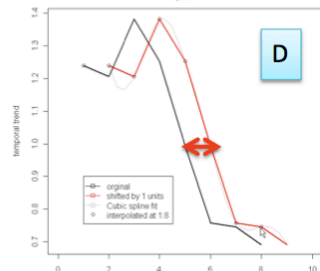
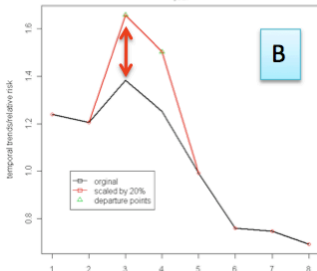
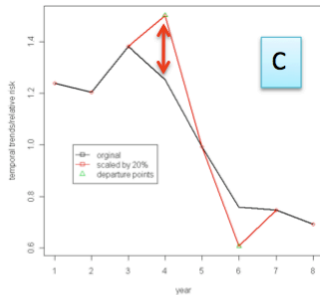
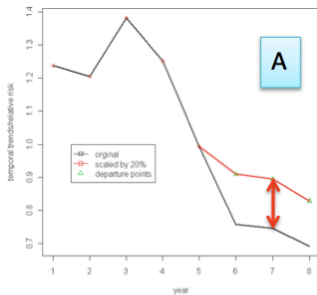
Summary: Income modelling

- The classification ability of the model depends on
 - the number of time points at which the local trend differs from the common trend;
 - the magnitude of the departure.
- The model performed considerably well in detecting trends with departure from the common at three or more time points and/or the departures are medium/large.
- This results are anticipated because the model is constructed to examine the difference of the overall pattern rather than depicting departures at individual time points.
- Although it is rare to have more than one replicate in practice, replicates reduce the uncertainty and hence provide better classification performance from the model.

- 1 Motivation
- 2 Mixture model
- 3 Application: Modelling Income
 - Data
 - Results
 - Summary
- 4 Simulation: Disease Surveillance**
 - Data and model
 - Results
 - Summary
- 5 Conclusion and Discussion

- In this simulation, the Poisson type of data is analyzed. It is more difficult to analyze since we do not have the control on the variance as in the Normal case;
- "True" unusual temporal patterns are constructed so that they are similar to what usually observed in practice.

Various "unusual scenarios"



Setting up the scene

- This simulation is based on a real dataset on county level asthma ambulatory incidence in Georgia, US (Lawson 2008).
- A separable space-time model was fitted to the incidence data:

$$y_{i,t} \sim \text{Poisson}(\theta_{i,t} \cdot E_{i,t})$$

$$\log(\theta_{i,t}) = \alpha + u_i + v_t + \epsilon_{i,t}$$

where u_i and v_t are the spatial and temporal components respectively and they are common to all areas.

- We then used the estimates for α , u_i and v_t to simulate count data for 90% of the areas (the "usual" cluster).
- Data for the remaining 10% of the areas (the "unusual" cluster) were simulated using α and u_i but v_t^* selected from one of the four "unusual" scenarios.

- For comparison, we also fitted a model, 2Cluster, that classifies areas according to their time trends into two clusters, namely, the "usual" and "unusual" clusters.
- This is the underlying true model.

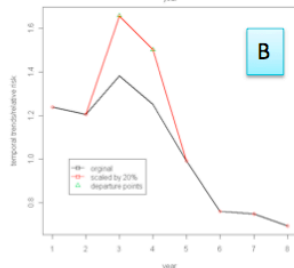
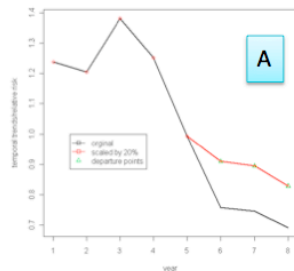
Results: AUC of scenarios A and B

Scenario A

Departure	Detect Model	2Cluster
20%	0.78	0.72
35%	0.87	0.66
50%	0.85	0.64

Scenario B

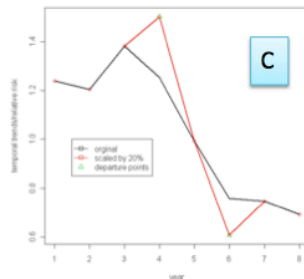
Departure	Detect Model	2Cluster
20%	0.77	0.76
35%	0.79	0.84
50%	0.80	0.98



Results: AUC of scenarios C and D

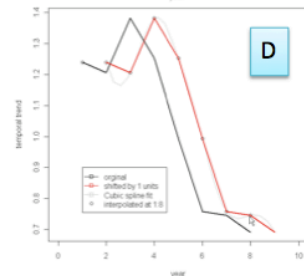
Scenario C

Departure	Detect Model	2Cluster
20%	0.86	0.79
35%	0.81	0.96
50%	0.78	1.00



Scenario D

Shift	Detect Model	2Cluster
0.25 tp	0.75	0.67
0.5 tp	0.76	0.68
1 tp	0.88	0.86



Summary of simulation

- The mixture detection model performed exceptionally well in all four scenarios, especially when the departures are large and/or occurring at more than one time point apart from Scenario C.
- For small departures, the detecting ability of the mixture model is slightly better than that of the "true" model with two trend clusters.

- 1 Motivation
- 2 Mixture model
- 3 Application: Modelling Income
 - Data
 - Results
 - Summary
- 4 Simulation: Disease Surveillance
 - Data and model
 - Results
 - Summary
- 5 Conclusion and Discussion

Conclusions and Discussions (I)

- The proposed mixture model fits a local trend using either the common trend model or an area-specific model and classifies areas based on the selection probability.
- The model performed reasonably well in both the income modelling application and the simulation.
- By construction the strength of this model is to detect unusual trends that have a different profile to that of the common trend, for example, higher peaks or the entire temporal profile shifted by some time units.
- Some preliminary results showed that this model may not be useful in detecting trends that fluctuate (without a distinctive structure) around the common trend (for example, as in Scenario C).

Conclusions and Discussions (II)

- A similar approach was proposed in the context of disease mapping by Abellan et al. (2008) who assumed a separable space-time model and classified areas with unusual trends by assessing the residuals. (Large residuals \equiv unusual)
- Comparisons of the two detection approaches: strengths and weaknesses.
- Applications to real data, e.g., to model rates of unemployment/crime trends/disease trends etc.

Acknowledgement

- This work is funded by the ESRC National Center for Research Methods through the BIAS II project.
- We are grateful to the Swedish Statistical Bureau for providing the income data.

ESRC National Centre for

Research
Methods



Refereces

- J.J. Abellan, S. Richardson, and N. Best. Use of space-time models to investigate the stability of patterns of disease. Environmental Health Perspectives, 116(8):1111-1119, 2008.
- B.P. Carlin and S. Chib. Bayesian model choice via Markov chain Monte Carlo methods. JRSSB, 57(3): 473-484, 1995.
- Lawson, A. B. (2008) Bayesian Disease Mapping: Hierarchical Modeling in Spatial Epidemiology, CRC Press (data are in Chapter 11).
- WinBUGS: Lunn DJ, et al. WinBUGSA Bayesian modelling framework: concepts, structure, and extensibility. Stat Comput 2000;10:32537. The program can be downloaded from <http://www.mrc-bsu.cam.ac.uk/bugs/>.

Defining the "true" unusual trends from the entire data

- To define a "true" unusual time trend, we need to decide
 - 1 the number of time points at which the local trend deviates from the common trend;
 - 2 whether it is unusual
 - 3 the amount it differing from the common trend at each time point;

